JOURNAL OF BACTERIOLOGY, Mar. 2009, p. 1480–1489 Vol. 191, No. 5
0021-9193/09/$08.00+0 doi:10.1128/JB.01343-08

Role of Conjugative Elements in the Evolution of the Multidrug-Resistant Pandemic Clone *Streptococcus pneumoniae*^{Spain23F} ST81$^{\nabla}$†

Nicholas J. Croucher,[1]* Danielle Walker,[1] Patricia Romero,[2] Nicola Lennard,[1] Gavin K. Paterson,[2] Nathalie C. Bason,[1] Andrea M. Mitchell,[2] Michael A. Quail,[1] Peter W. Andrew,[3] Julian Parkhill,[1] Stephen D. Bentley,[1] and Tim J. Mitchell[2]

The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, Cambridgeshire CB10 1SA, United Kingdom[1]*; Division of Infection and Immunity, Glasgow Biomedical Research Centre, University of Glasgow, 120 University Place, Glasgow G12 8TA, United Kingdom*[2]*; and Department of Infection, Immunity and Inflammation, Medical Sciences Building, University of Leicester, Leicester LE1 9HN, United Kingdom*[3]

Received 26 September 2008/Accepted 16 December 2008

***Streptococcus pneumoniae* is a human commensal and pathogen able to cause a variety of diseases that annually result in over a million deaths worldwide. The *S. pneumoniae*^{Spain23F} sequence type 81 lineage was among the first recognized pandemic clones and was responsible for almost 40% of penicillin-resistant pneumococcal infections in the United States in the late 1990s. Analysis of the chromosome sequence of a representative strain, and comparison with other available genomes, indicates roles for integrative and conjugative elements in the evolution of pneumococci and, more particularly, the emergence of the multidrug-resistant Spain 23F ST81 lineage. A number of recently acquired loci within the chromosome appear to encode proteins involved in the production of, or immunity to, antimicrobial compounds, which may contribute to the proficiency of this strain at nasopharyngeal colonization. However, further sequencing of other pandemic clones will be required to establish whether there are any general attributes shared by these strains that are responsible for their international success.**

Streptococcus pneumoniae (the pneumococcus) is a human commensal and pathogen that represents a major cause of otitis media, pneumonia, and meningitis (8). Worldwide, pneumococcal disease is thought to be responsible for over a million fatalities annually, including more than 800,000 deaths in children under 5 years of age living in developing countries (64). While the introduction of the heptavalent polysaccharide conjugate vaccine (PCV7) has dramatically reduced the incidence of pneumococcal disease in some areas (37), limited serotype coverage, strain replacement, and capsule switching have resulted in a smaller, and decreasing, impact in other communities (66).

S. pneumoniae is a naturally competent, genetically diverse species, with less than half of the pan-genome conserved between all strains thus far studied (33). The pneumococcal population is normally confined to the human nasopharynx, with rates of asymptomatic carriage varying with demographics, region, and season: surveys of colonization in healthy children generally estimate between 20 and 97% of younger individuals carry pneumococci (9, 32), with levels falling with age. Epidemiological data and animal models of infection indicate that strains exhibit differing propensities for causing invasive disease (13, 29, 63). The invasive disease potential odds ratio, which takes into account the relative frequencies of invasive disease and asymptomatic carriage observed in the human population, varies 80-fold between serotypes (13). However, the functional genetic variation to which this differing ability to cause disease is attributable remains largely unknown. Genome sequencing efforts have mainly focused on clinical pneumococcal isolates; the complete genomes of two highly invasive strains, TIGR4 (70) and D39 (38) (and the laboratory-adapted D39 derivative R6) (34), have been published, along with draft sequences for serotype 19F strain G54 (26) and eight clinical isolates from a hospital in Pittsburgh (65). However, in order to understand the bacterial population structure, and the reasons underlying the variation in pathogenicity, genomic studies of strains that only rarely invade past the mucosal surfaces are required.

S. pneumoniae^{Spain23F} sequence type 81 (ST81) was one of the first pandemic penicillin-resistant clones identified (47). Initially characterized among isolates from Spain in the 1980s, it spread globally, and by the late 1990s it was estimated to constitute almost 40% of penicillin-resistant disease isolates in the United States (21). The clone is also resistant to chloramphenicol and tetracycline and is one of those most frequently associated with the emergence of fluoroquinolone and macrolide resistance (58, 60). This lack of susceptibility to the major classes of antimicrobial chemotherapies used to treat pneumococcal infections has undoubtedly aided the spread of the strain and lead to the inclusion of the 23F serotype in PCV7. This has resulted in a reduction in the prevalence of the *S. pneumoniae*^{Spain23F} ST81 clone in some regions (35). However, the lineage has undergone capsule switching to alternative capsule types on at least three occasions (from serotype 23F to 14 (36), 19A (20), and 19F (19), suggesting it is liable to eventually evade the vaccine. Despite its high carriage preva-

* Corresponding author. Mailing address: Sulston Laboratories, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, Cambridgeshire CB10 1SA, United Kingdom. Phone: 44 1223 834244. Fax: 44 1223 494919. E-mail: nc3@sanger.ac.uk.

† Supplemental material for this article may be found at http://jb.asm.org/.

$^{\nabla}$ Published ahead of print on 29 December 2008.

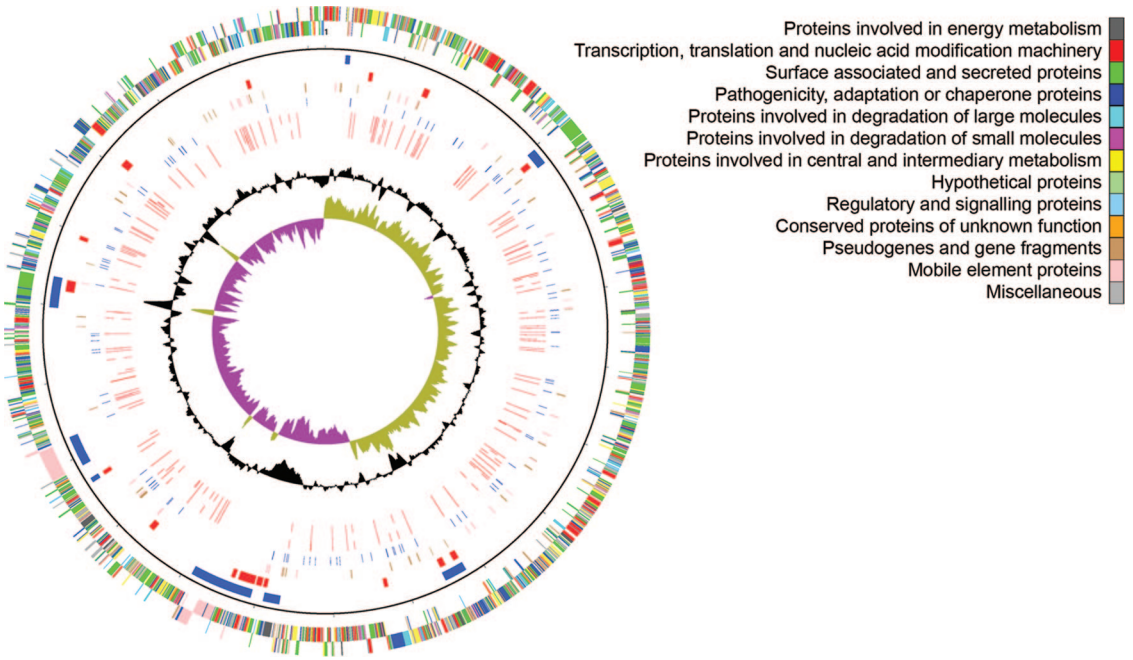

FIG. 1. Circular diagram representing the *S. pneumoniae* ATCC 700669 chromosome, arranged to have the origin of replication at the top, as indicated by the GC deviation (innermost graph). The outer rings show the arrangement of coding sequences on the two strands of the genome, colored according to annotated function (see the key). The first inner ring indicates the major variable regions as blue blocks: moving clockwise from the origin of replication, these represent the prophage remnant, *cps* locus, PPI-1, the Na^+-dependent ATP synthase island, ICE*Sp*23FST81, the *pclA* gene cluster, ϕMM1-2008, and the *psrP* gene cluster. The red blocks demarcate loci identified as having atypical nucleotide composition by the Alien Hunter algorithm (71). Moving inward, the rings show the position of IS elements (pink if intact, brown if pseudogene), RUP repeats (blue), and BOX A, B, and C repeat modules (red), respectively. The black graph indicates sequence GC content.

lence (22), it has a low propensity for causing invasive disease (67) (odds ratio of 0.4 [13]), suggesting its intercontinental distribution has been facilitated by adaptations to colonization of, and survival within, the human nasopharynx. Here we report our analysis of the complete genome of *S. pneumoniae* ATCC 700669, a member of the serotype 23F ST81 lineage that was isolated in a hospital in Barcelona in 1984 (18).

MATERIALS AND METHODS

Genome sequencing. A shotgun sequence with ~8-fold genome coverage was achieved through sequencing of pUC clones with 1.4- to 2.8-kb inserts and pSMART clones with 8- to 12-kb inserts using a BigDye terminator sequencing kit (Applied Biosystems) and Applied Biosystems 3700 sequencers. Sequences from 30- to 40-kb pEpiFOS-5 fosmid clones and 12- to 23-kb pBACe3.6 BAC clones were used to scaffold contigs and bridge repeats. The sequence was finished according to standard criteria (55). Sequence assembly, visualization, and finishing were performed by using PHRAP (www.phrap.org) and Gap4 (10). All repeat sequences were independently verified.

Annotation and genome comparisons. Coding sequences were initially identified by using Glimmer3 (24) and then manually curated using Frameplot (7) and Artemis (62). All genes were annotated in Artemis using standard criteria (6). Genome comparisons were visualized in the Artemis comparison tool (16). Sequence clustering and analysis was performed by using ClustalX 2.0 (39) and MEGA4 (69).

Nucleotide sequence accession numbers. The sequence and annotation of *S. pneumoniae* ATCC 700669 has been deposited in the EMBL database under the accession number FM211187. Other sequences referred to in the present study are also deposited in the public databases with the following accession numbers: *S. pneumoniae* TIGR4 (AE005672), D39 (CP000410), R6 (AE007317), G54 (CP001015), and CGSP14 (NC010582) and the eight draft sequences of pneumococcal clinical isolates from Pittsburgh (3-BS71, AAZZ01000001-AAZZ01000024; 6-BS73, ABAA01000001-ABAA01000038; 9-BS68, ABAB01000001-ABAB01000061; 11-BS70, ABAC01000001-ABAC01000025; 14-BS69, ABAD01000001-ABAD01000049; 18-BS74, ABAE01000001-ABAE01000028; 19-BS75, ABAF01000001-ABAF01000030; and 23-BS72, ABAG01000001-ABAG01000032).

RESULTS AND DISCUSSION

Genome structure. The genome of *S. pneumoniae* ATCC 700669 is a circular chromosome of 2,221,315 bp (39.49% GC content; Fig. 1) containing four rRNA operons and 58 tRNA genes (all but 12 of which are adjacent to rRNA genes). An unusual asymmetry in the GC skew of the chromosome, resulting from the recent integration of a prophage and an integrative and conjugative element (ICE) into the same replichore, is evident. There are 2,135 predicted coding sequences (CDS), 144 (6.7%) of which appear to be pseudogenes. In common with other *Firmicutes*, there is a strong coding bias, with 76% of the CDS on the leading strand. Overall, 197 (9.2%) of the predicted CDS in ATCC 700669 are not present in TIGR4 or D39, the majority of these being present on an ICE or prophage-derived sequences.

In total, 4.4% of the ATCC 700669 genome is composed of interspersed repeats. The sequence includes 125 BOX repeats, complex modular elements consisting of various combinations of A, B, and C components (43), and 115 RUP repeats, which are simpler 107-bp elements hypothesized to be mobilized in *trans* by the IS*630* transposase (51). Seventy-nine insertion sequence (IS) elements can be identified, of which 73% appear to be no longer functional. Twenty-one of the IS insertions are not present in the TIGR4 or D39 genomes, the majority of which are due to IS*1167* or IS*1167A*-type elements. Sequence-

based clustering of the IS elements supports the explanation that these transposons have undergone a recent expansion in the 23F ST81 lineage (see Fig. S1 in the supplemental material), in accordance with the previously proposed hypothesis that IS isotypes spread through short bursts of transposition (70).

Of the annotated genes, 29% are predicted to encode surface exposed or secreted proteins. These include 18 phosphotransferase system (PTS) transporters, 17 of which are shared with both TIGR4 and D39, which have four and three PTS transporters not present in ATCC 700669, respectively. The one system unique to ATCC 700669 (SPN23F18210-18250) forms part of a ~10-kb insertion that also includes a putative choline sulfatase. Given the importance of choline to pneumococcal metabolism and pathogenesis, this transporter could represent a novel means of acquiring this nutrient from the host environment. *S. pneumoniae* genomes also encode many ATP-binding cassette (ABC) transporter components, and there are two such systems within the ATCC 700669 genome that are absent from both TIGR4 and D39. One of these, present on a ~4-kb insertion along with two putative secreted peptides (SPN23F07060-07090), is similar to systems present in *S. pyogenes* MGAS6180, *S. pyogenes* MGAS10270, and *S. sanguinis* SK36.

Genes implicated in pathogenesis. Genome-wide screens in *S. pneumoniae* have identified a large number of genes required for full virulence in the mouse model of pneumococcal disease (31, 40, 54, 59), and variation in their sequences may lie behind the differences in invasiveness seen in epidemiological studies. One such "virulence factor" is PspA, a highly variable choline-binding surface protein that protects *S. pneumoniae* against the bactericidal effects of apolactoferrin, as well as inhibiting complement-mediated opsonization of the pneumococcus (5). It has been shown that the *pspA* gene of a serotype 23F strain contained a frameshift mutation that truncated the encoded protein to a secreted, rather than surface-associated, form (44). A similar frameshift mutation, caused by variation in the length of the same polyadenosine tract, is observed in this genome. Furthermore, ATCC 700669 lacks the *zmpC* gene, which encodes a metalloprotease that cleaves and activates human matrix metalloprotease 9 and has been implicated in the pulmonary invasion process (52). However, like *S. pneumoniae* G54, it encodes an additional surface-exposed zinc metalloprotease, *zmpD* (15), adjacent to the immunoglobulin A protease gene *zmpA*. Notably absent from the ATCC 700669 genome are two more surface-exposed proteins, choline-binding protein *cbpC* and histidine triad protein *phtB*, which appear to be required for full virulence of TIGR4 in the mouse model, based on signature-tagged mutagenesis data (31).

There are two large islands present in *S. pneumoniae* ATCC 700669 that are shared with TIGR4 but absent from D39, both of which were found to be overrepresented in invasive strains in a comparative genomic hybridization study of clinical isolates (49). One, region of diversity 8a, is a ~25-kb island encoding an Na^+-dependent ATP synthase. The second is the *psrP* gene cluster, encoding a large adhesin similar to the *S. gordonii* platelet-binding protein GspB, along with associated glycosylases (61). Although ATCC 700669 lacks both identified pneumococcal pilus synthesis gene clusters (2, 3), it does possess the large surface-anchored collagen-like protein PclA (56), which contains G5 domains implicated in biofilm formation (4). Absent from TIGR4, the D39 orthologue is ~44% longer due to an expansion in the internal repetitive region of the protein.

Prophage. Prophage are bacteriophage genomes integrated into the host bacterial chromosome, which have been shown to be responsible for the transduction of virulence factors between pathogens on numerous occasions (72). Two closely related temperate pneumophage, ϕMM1, from a serotype 23F ST81 strain (50), and ϕMM1-1998, from a serotype 24 strain (41), have been previously sequenced. Although the *S. pneumoniae* G54 and D39 genomes are devoid of prophage, a 10.5-kb phage remnant can be found between the *eno* and *rexB* genes in the chromosome of TIGR4 (50). The ATCC 700669 genome contains an intact 39.1-kb prophage, ϕMM1-2008, as well as a smaller prophage remnant. ϕMM1-2008 is more than 97% identical, at the nucleotide sequence level, to both ϕMM1 and ϕMM1-1998. Since ϕMM1 and ϕMM1-2008 are from hosts of the same serotype and sequence type, they are probably descended from the same insertion event, while the host of ϕMM1-1998 is a penicillin-sensitive serotype 24 strain; hence, this prophage is likely to be the result of a different infection. All three phage are present in the same locus, between a pyridine nucleotide-disulfide oxidoreductase gene and a CDS of unknown function. The exact insertion site of the phage appears to be within the 3′ region of the downstream hypothetical gene; duplication of this 15-bp *att* sequence (30) upon integration maintains the full-length target gene sequence and generates the tandem repeats either side of the prophage. Acquisition of ϕMM1-1998 by pneumococcal hosts was correlated with increased adhesion to eukaryotic cells (41), implying prophage genes may potentially play a role in colonization or pathogenesis.

The 6.4-kb prophage remnant is flanked by genes encoding a putative cytidine deaminase and a deoxyuridine 5′ triphosphate nucleotidohydrolase. Along with CDS of phage origin, including integrase and amidase pseudogenes, the prophage appears to carry "cargo" genes that have been retained as the replicative machinery of the virus has degenerated. One of these CDS encodes a type I restriction endonuclease domain and appears to be a member of a family of uncharacterized genes found in a range of bacterial species. Another is an addiction system toxin gene, which may have inhibited the clearance of the remainder of the prophage from the genome. Although a cognate antitoxin cannot be reliably identified, the overlapping upstream gene is a good candidate, as alignments with intact prophage indicate these CDS have been acquired as a pair. A complete 37.5-kb long prophage can be found at this locus in the draft genome of *S. pneumoniae* 18-BS74 (65), suggesting it may be a common target insertion site for temperate pneumophage.

ICE. ICEs are autonomously mobile genetic entities that are transmitted between bacteria through conjugative transfer in a circularized form and insert into, and excise from, host DNA via site-specific recombination (14). This grouping includes both conjugative transposons and integrative plasmids that share this common mechanism of horizontal transfer. The ~81-kb ICE of ATCC 700669, ICE*Sp*23FST81, is a composite element comprising a Tn*916*-like component inserted into a Tn*5252*-like transposon, with the latter consequently being

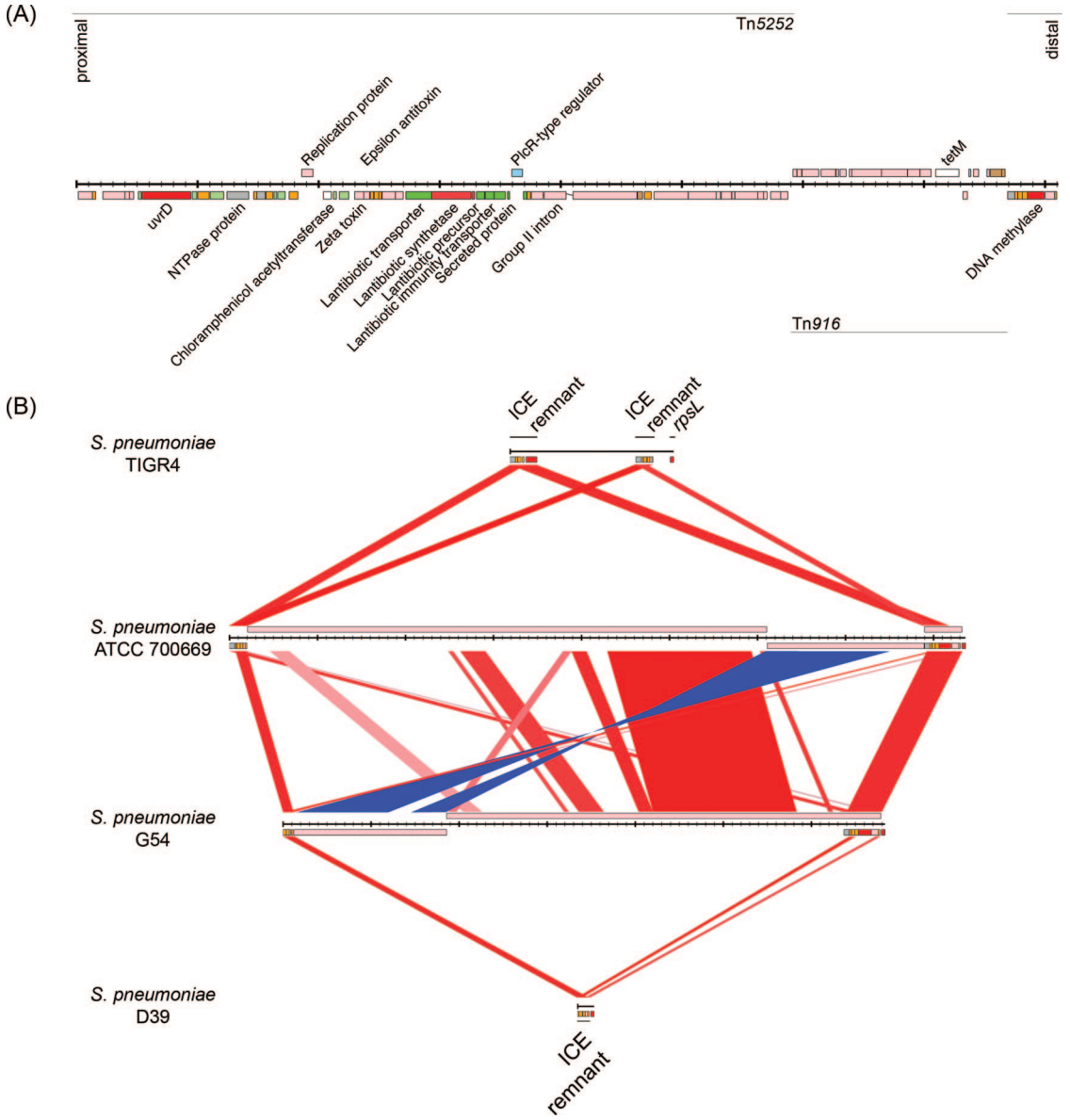

FIG. 2. (A) Representation of ICE*Sp*23FST81. "Cargo" genes are colored according to the scheme detailed in Fig. 1 and are labeled with their putative function, where one can be assigned. The division of the ICE into Tn*5252* and Tn*916*-type elements is indicated by the bars at the top and bottom of the diagram. (B) Comparison of the ICE insertion sites in the ATCC 700669 and G54 genomes with the corresponding loci in the TIGR4 and D39 genomes, which lack intact ICE. The ICE are represented as pink blocks: Tn*5252*-type elements are represented by blocks above the scale line, while Tn*916*-type elements are represented by blocks below the scale line. Red bands indicate BLAST matches between genomes in the same direction, whereas blue bands indicate matches in opposite directions. The intensity of the band represents the strength of the match. The region of each genome shown is bounded at the 5′ end by ICE-derived sequences and at the 3′ end by *rplL*, which streptococcal ICEs frequently insert directly upstream of. The ICE remnants in the TIGR4 and D39 genomes, apparently derived from the distal region of the Tn*5252*-type element, are marked. The alignment shows that there is a remnant in D39 directly adjacent to *rplL*, at the point where the elements usually insert, but the two remnants in the TIGR4 genome are further removed upstream.

split into a larger proximal region and a smaller distal region (Fig. 2A). This combination of conjugative transposons is common in streptococcal ICEs, although the variation observed in the arrangement of these two elements implies it has arisen independently on a number of occasions (Fig. 3). This suggests a potential symbiotic advantage between these different transposon types, perhaps resulting from a synergistic combination of the two sets of conjugative machinery or "cargo" genes. Flanked by a 16-bp tandem duplication, ICE*Sp*23FST81 is inserted near the 3′ end of *rplL*, in the same position as that of *S. pneumoniae* G54 and many other streptococcal ICEs, although Tn*5253*, the partially sequenced ICE of *S. pneumoniae* BM6001 (1), appears to have integrated elsewhere.

Shortly upstream of *rplL* in the TIGR4 and D39 genomes are ~1.8-kb long ICE remnants that are >80% identical, at the nucleotide level, to the distal Tn*5252*-type region of ICE*Sp*23FST81 (Fig. 2B). Similar remnants are also seen in the ATCC 700669 and G54 genomes immediately upstream of

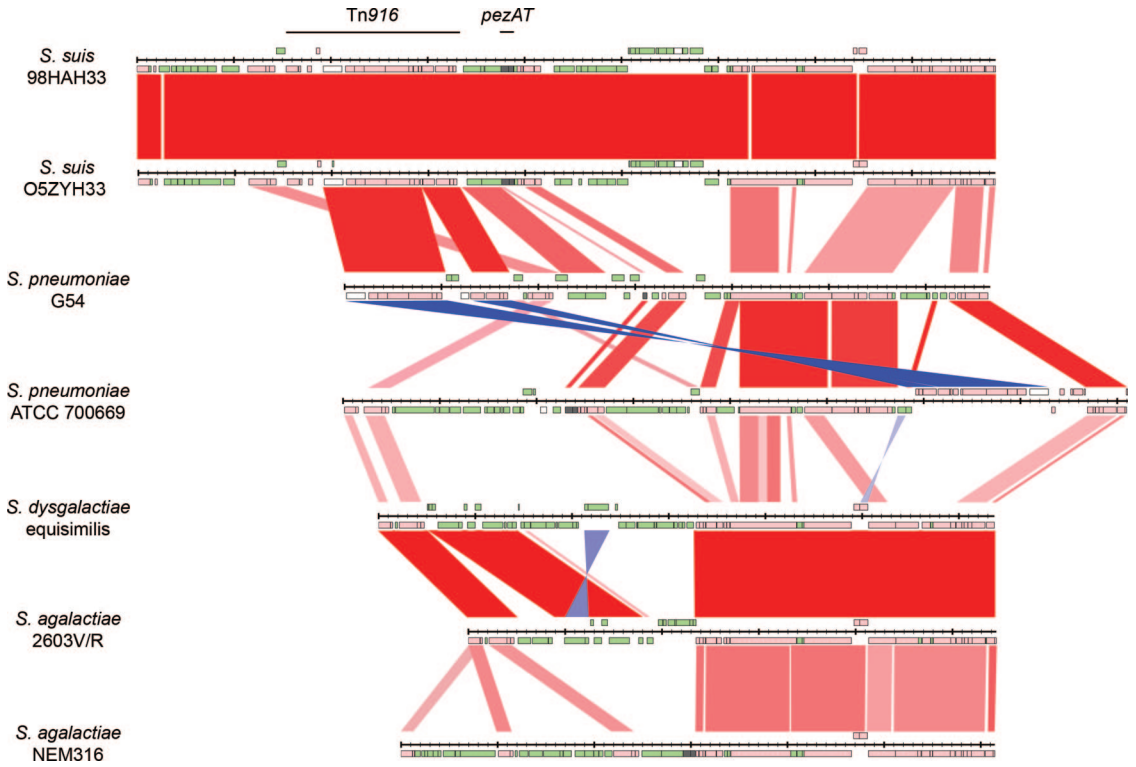

FIG. 3. Comparison of streptococcal integrative and conjugative elements. Genes likely to be part of the conjugative machinery of the element, on the basis of conservation or functional assignment, are colored pink. Zeta toxin-epsilon antitoxin systems are colored gray, antibiotic resistance genes are white, and other "cargo" genes are green. Bands are colored as in Fig. 2, with the intensity of the color indicating the strength of the BLAST match. Unlike the *S. suis* and *S. pneumoniae* elements, the *S. agalactiae* and *S. dysgalactiae* Tn*5252*-like elements do not contain a Tn*916*-type component, which carries the *tetM* gene responsible for tetracycline resistance in the other ICE.

their ICE insertions. A second, larger, ICE remnant is also evident in the TIGR4 genome, ~15 kb upstream of *rplL*; this includes a cytosine methyltransferase gene very similar to homologues in the distal region of ICE*Sp*23FST81, on Tn*5253* and in the ϕMM1 phage (85%, 85% and 45% protein sequence identity, respectively). The presence of this gene on conjugative transposons and prophage suggests it may aid the horizontal transfer of both between pneumococci, perhaps through methylating DNA prior to transfer between cells and hence avoiding the recipient's restriction systems.

ICE*Sp*23FST81 is clinically important due to its genetic "cargo." The Tn*916*-type component carries a *tetM* gene, responsible for the strain's tetracycline resistance. A similar Tn*916* element, also carrying the *mef*(A) macrolide resistance gene, has been detected in the gammaproteobacterial commensal and emerging nosocomial pathogen *Acinetobacter junii* (53). The ~1.2-kb flanking sequences that were determined on either side of the *A. junii* Tn*916* transposon are 99% identical, at the nucleotide level, to the Tn*5252*-type sequences surrounding the Tn*916* transposon on ICE*Sp*23FST81, suggesting these composite elements can transfer between very distantly related bacteria. The Tn*5252*-like component carries a gene for chloramphenicol acetyltransferase, which appears to have been acquired through wholesale integration of the pC194 plasmid (73) originally identified in chloramphenicol-resistant *Staphylococcus aureus*.

One of the "cargo" genes found toward the 5′ end of the element is a *uvrD* helicase gene, with the closest sequenced homologue being that of the deltaproteobacterium *Geobacter lovleyi* (26% protein sequence identity). A different *uvrD* gene, with a dissimilar sequence (15% protein sequence identity), is present at the equivalent position on the G54 ICE. Streptococci lack an SOS response (27), so consequently *uvrD* is absent from the *S. pneumoniae* core genome, but horizontal acquisition of this gene could potentially reconstitute the nucleotide excision repair pathway if it were able to act in concert with the *uvrABC* genes shared by all pneumococci. This pathway is important in the repair of peroxidative damage to DNA (46); therefore, given that *S. pneumoniae* is catalase negative and produces hydrogen peroxide which can function as an antimicrobial (57), the gain of the *uvrD* gene may increase the tolerance of ATCC 700669 to reactive oxygen species and hence aid nasopharyngeal colonization, while also resulting in the ICE maintaining its sequence integrity within the host. The other major branch of the SOS response, also absent from the core genome of pneumococci, is mutagenic lesion repair, which requires a reduction in DNA polymerase III replication fidelity caused by an interaction with the UmuC-UmuD complex. Correspondingly, Tn*5252* carries a *umuCD*-containing operon that was demonstrated to increase the UV tolerance of the host bacterium (48), suggesting ICE-carried genes can functionally restore at least one aspect of the SOS response.

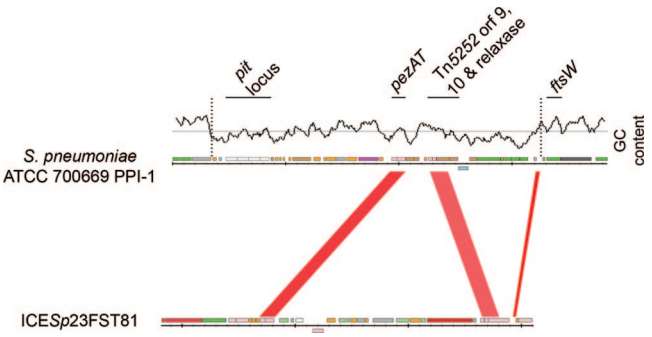

FIG. 4. Comparison of PPI-1 of ATCC 700669 with ICE*Sp*23 FST81 found in the same strain. Three regions of apparently conserved similarity are observed: the zeta toxin-epsilon antitoxin system, a group of Tn*5252* conjugative transfer genes (encoding ORF 9, ORF10, and relaxase proteins), and a short stretch of DNA adjacent that forms the 3′ border of PPI-1, shortly upstream of the cell division gene *ftsW* in the pneumococcal chromosome. The GC content of this region is shown, with the line across the graph indicating the average for the region (33.71% GC). The vertical dotted lines on the graph delimit the extent of PPI-1.

An ICE-derived genomic island. Genes characteristic of streptococcal ICE are also found in another region of the genome, within the putative pathogenicity island 1 (PPI-1), as described by Brown et al. (12). In *S. pneumoniae* ATCC 700669, this ~30-kb region, as defined by its low GC content (SPN23F09511-09860; Fig. 4), contains the *pezAT* epsilon toxin-zeta antitoxin system, found on ICE*Sp*23FST81, as well as related elements in *S. suis* and *S. agalactiae* strains, and a cluster of three Tn*5252* conjugative machinery genes, including a relaxase and a MobC-domain protein. At the 3′ end, coinciding with the edge of the low GC region, there is a further ~200-bp region of similarity (>90% identity at the nucleotide level) with ICE*Sp*23FST81, which is also shared with Tn*5253* and the ICE of the newly sequenced *S. pneumoniae* CGSP14 strain (accession number NC010582). A site-specific recombinase, similar to one found in *S. suis* ICEs, is found adjacent to this sequence in some pneumococcal strains, such as *S. pneumoniae* 14-BS69. Hence, it appears likely that this island originated as an ICE insertion that has subsequently degenerated, with the loss of genes required for the element's autonomous mobility. The gene clusters located between these ICE-like regions are very different in many of the strains for which genome data are available (Fig. 5), suggesting this locus may be able to diversify through exchange of sequences with ICE via homologous recombination in the shared regions.

A further source of variation appears to be the extent to which the conjugative machinery at this locus has degenerated (Fig. 5), with *S. pneumoniae* 18-BS74 missing all of the ICE-like regions of the island. In contrast, the 5′ end of PPI-1, containing the *pit* iron transporter operon crucial for virulence of *S. pneumoniae* (11), is conserved among all strains. However, it seems likely the *pit* genes were acquired as part of the original ICE insertion, since they also lie within the low-GC-content region and, despite being ubiquitous among sequenced pneumococci, appear to be absent from other streptococci (11).

The ability to exchange sequences with conjugative elements may play a role in allowing this locus to rapidly evolve in response to changing selection pressures. In strains that have retained the variable region, CDS similar to genes encoding daunorubicin efflux transporters, inhibitor-resistant methionyl tRNA synthetases, macrolide efflux pumps, and apparently

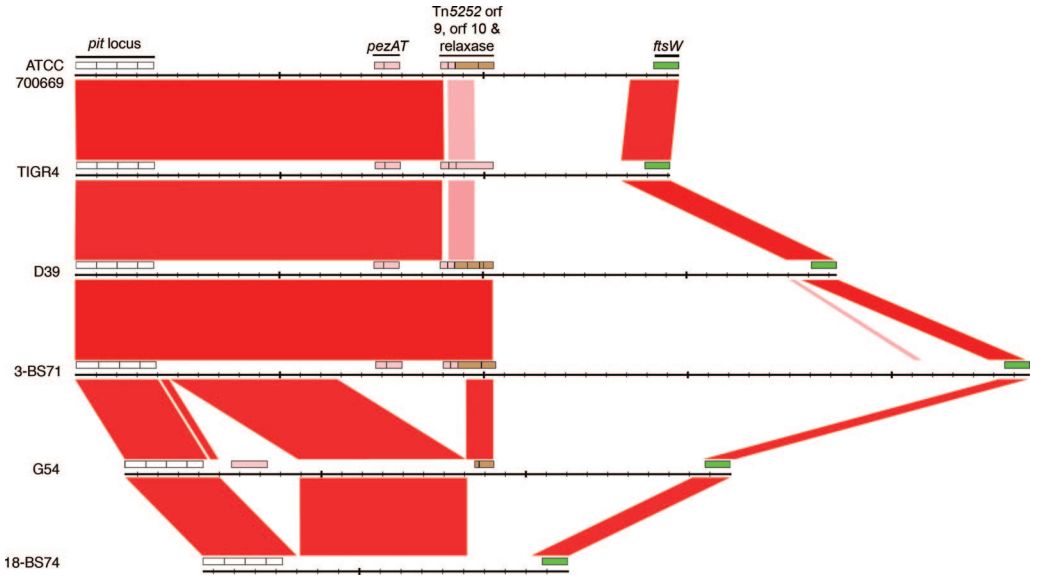

FIG. 5. Alignment of PPI-1 from complete and draft pneumococcal genome data. Variation in the regions intervening between the relaxase and 3′ end of PPI-1 (these boundaries both have sequence homology with Tn*5252*-type ICE) appears to be due to horizontal gene transfer, indicating that conjugative elements may contribute to the diversity found within this island through homologous recombination-mediated exchange. There is also variation putatively resulting from degeneration of the original conjugative element insertion: the PPI-1 of *S. pneumoniae* G54 has lost the *pezAT* toxin-antitoxin system and much of the cluster of conjugative transfer genes (although it has retained a CDS encoding a MobA-domain protein, which are typically associated with autonomously mobile elements, indicated in pink), while that of 18-BS74 appears to have lost all vestiges of the original element's conjugative machinery.

truncated aminoglycoside phospho- and acetyltransferases are found. Furthermore, two of the sequenced Pittsburgh disease isolates (9-BS69 and 14-BS69) seem to carry chloramphenicol acetyltransferase genes on the island. This is at least the fourth documented case in which Tn*5252*-type elements appear to have contributed to the transfer of chloramphenicol acetyltransferase genes to pneumococci. Tn*5253* of *S. pneumoniae* BM6001 and ICE*Sp*23FST81, both intact, potentially mobile, elements, carry chloramphenicol acetyltransferase genes at different positions. The PPI-1 locus and the IQ complex of *S. pneumoniae* 529, a genomic island containing Tn*5252*- and Tn*916*-like fragments, along with chloramphenicol acetyltransferase and macrolide resistance genes (45), both appear to be fragments of conjugative transposons that have been integrated into the chromosome, suggesting that exchange of DNA between ICE and the pneumococcal genome is likely to be an important mechanism in the dissemination of antibiotic resistance throughout the *S. pneumoniae* population.

In *S. pneumoniae* ATCC 700669, PPI-1 contains an apparently incomplete lantibiotic synthesis operon. Lantibiotics (lanthionine-containing antibiotics) are small, secreted cyclic peptides containing lanthionine rings formed by the stereospecific intramolecular addition of cysteine to dehydrated serine or threonine residues (74). They frequently function as bacteriocins, with different types hypothesized to operate through inhibiting peptidoglycan transglycosylation or forming pores in cell membranes, but are also known to act as biosurfactants and phospholipase A2 inhibitors (74). The gene cluster present in the ATCC 700669 PPI-1 lacks a structural prepeptide gene but retains the CDS necessary for immunity. Comparison with the same locus in the serogroup 23 Pittsburgh isolate genome reveals a ~6-kb deletion in the ATCC 700669 gene cluster (Fig. 6B). The sequence absent in ATCC 700669 is flanked by thymidine dinucleotides and encodes two lantibiotic structural genes and fragments of two dehydratases. The putative product of the intact locus is a novel dimeric lantibiotic (Fig. 6A) likely to be similar to those produced by *Bacillus* and *Lactococcus* species (74).

Pneumocidins. The *blp* locus has been demonstrated to be important in intraspecific competition between pneumococci in a mouse model of colonization, and the ability of certain *S. pneumoniae* strains to inhibit the growth of others has been attributed to sequence variation within this gene cluster (23, 42). In ATCC 700669, the bacteriocin-producing *blp* locus has undergone a rearrangement relative to that in TIGR4 (see Fig. S2 in the supplemental material), resulting in the deletion of *blpM* and *blpN*, both of which are required for *blp*-encoded bactericidal activity (23, 42). Given the high level of nasopharyngeal carriage of Spain 23F ST81 strains, it seems likely that loci elsewhere in the genome are able to compensate for the loss of this important pneumocidin. Consistent with this suggestion, bacteriocin production by Spain 23F ST81 strains has been observed to inhibit the growth of a larger number of indicator strains than a penicillin-sensitive serotype 23F isolate (42). In addition to the defunct operon in the ATCC 700669 PPI-1, ICE*Sp*23FST81 itself carries an intact lantibiotic synthesis gene cluster. The structural gene appears to be a novel group II lantibiotic, most closely related to mersacidin and lichenicidin (SPN23F12701; Fig. 6C), on the basis of its probable ring structure, diglycine cleavage motif, and net neutral charge. Adjacent to this operon is the probable transcriptional regulator, which is similar to *plcR* of *Bacillus* species. Such transcription factors can form a minimal quorum-sensing system in conjunction with peptide autoinducers (68) and, in keeping with such a role, a small secreted peptide is found between the lantibiotic synthesis genes and the *plcR* homologue.

A similar *plcR* homologue-secreted peptide combination is present on the G54 ICE and also adjacent to a different ~10-kb putative lantibiotic synthesis gene cluster (SPN23F19690-19790) conserved in the chromosomes of ATCC 700669, TIGR4 and D39, next to the *rpoBC* genes. This locus also appears to be a recent addition to the genome on the basis of its atypical nucleotide composition (71) and absence from other pneumococcal sequences (e.g., *S. pneumoniae* OXC141 and INV104 [www.sanger.ac.uk/Projects/S_pneumoniae/]). In addition, shortly upstream of the ICE*Sp*23FST81 insertion site lies another region (SPN23F12290-12330) in these three genomes that appears to have been recently horizontally acquired, on the basis of its nucleotide composition and flanking tandem repeats, which contains the remains of a set of lantibiotic processing machinery. Hence, the range of bacteriocins that individual pneumococcal strains are able to produce appears to change over time as new gene clusters are gained and old ones degenerate. If these antimicrobial peptides are associated with specific quorum-sensing systems, this could further increase the scope for competition between strains, since differences in signaling and regulation of the bacteriocins would also vary between different genotypes. This is likely to result in strains having a variable secretome that could strongly influence intra- and interspecific competition within the nasopharynx.

Concluding remarks. The international spread of *S. pneumoniae*$^{\text{Spain23F}}$ ST81 indicates the clone is highly effective at nasopharyngeal colonization, and the genome analysis presented here offers some possible explanations. The loss of the *blpMN* pneumocidin and PPI-1-encoded lantibiotic in ATCC 700669 suggests they may have become redundant during the evolution of the Spain 23F ST81 lineage. In both cases, the structural genes were lost, while those required for immunity were retained, ensuring the strain remained nonsusceptible to these compounds secreted by its competitors. A possible explanation is that the pneumocidin on the ICE may have superceded other bacteriocins produced by the strain and assisted nasopharyngeal colonization, and hence the spread, of the Spain 23F ST81 clone. Both this lantibiotic synthesis gene cluster and another elsewhere in the chromosome are found adjacent to genes encoding *plcR*-type transcriptional regulators and secreted peptides. If such quorum-sensing systems are involved in the regulation of bacteriocin production, this would add further layers of complexity onto the intercellular signaling already known to occur in *S. pneumoniae* (the two previously characterized pneumocidins, BlpMN and CibAB, are regulated by the pheromones BlpC and ComC, respectively) (17, 25). Such complexity may be necessary, in an organism with such a variable chromosome, to distinguish clonal relatives from other related strains. Furthermore, cell-density-dependent regulation of the ICE-encoded lantibiotic synthesis gene cluster may be advantageous for the mobile element itself, since suppressing production of the antimicrobial compound

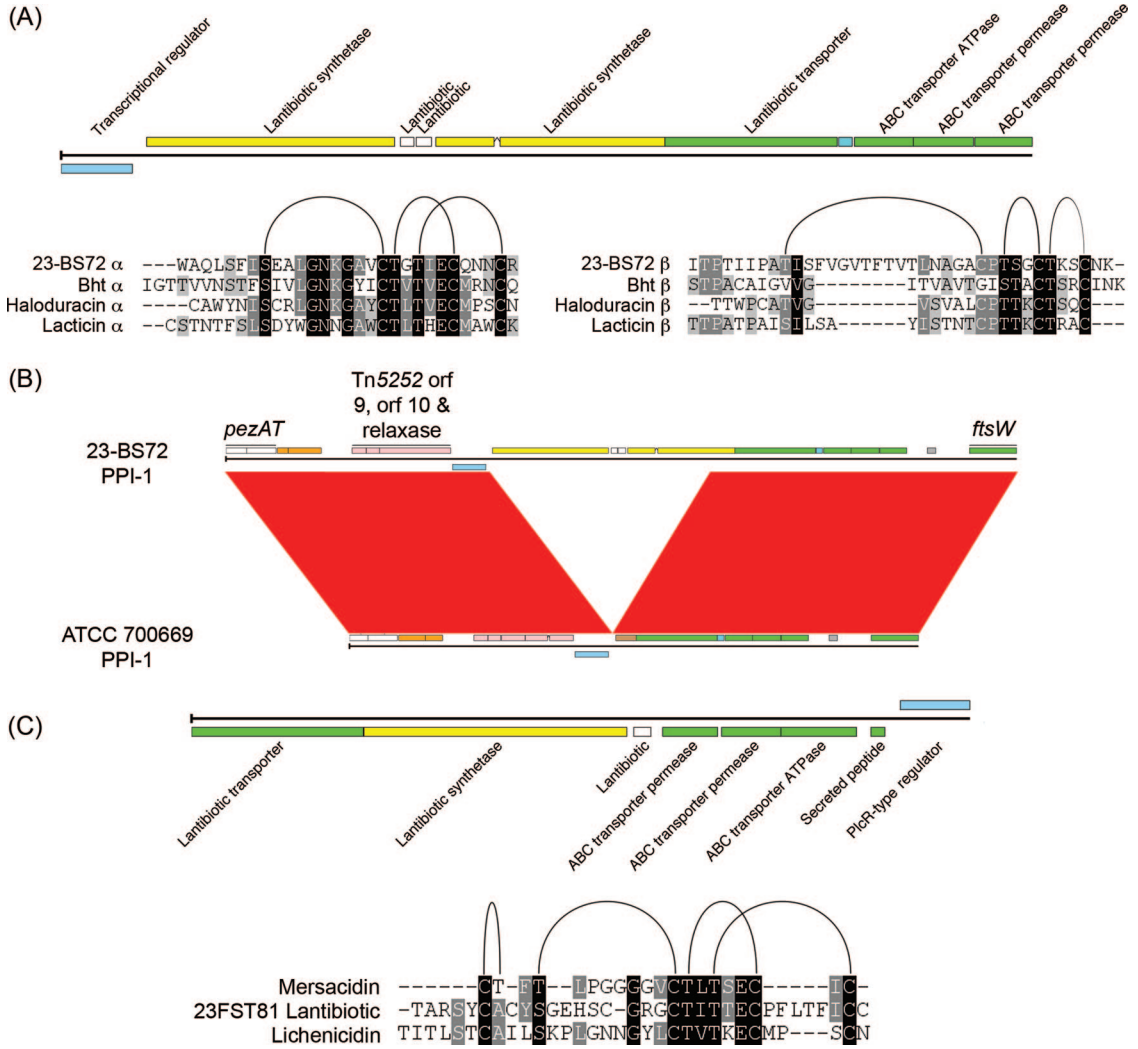

FIG. 6. Lantibiotic synthesis gene clusters and predicted structures. The color scheme is the same as that used in Fig. 1. (A) Gene cluster encoded on the PPI-1 of 23-BS72. Each structural gene appears to be associated with its own synthetase. The ABC transporter is thought to be responsible for the self-immunity of the producer. The likely sequences of the mature peptides are shown aligned with other dimeric lantibiotics, with the lanthionine rings (formed by the dehydration of cysteine and serine or threonine residues) conserved between haloduracin and lacticin 3147 indicated by the curved lines above the alignment. The conservation of the functionally most important residues with these other bacteriocins indicates the 23-BS72 lantibiotic is likely to be bactericidal. (B) Alignment of the PPI-1 of *S. pneumoniae* 23-BS72 and ATCC 700669, revealing a ~6-kb deletion in the gene cluster of the latter. (C) Gene cluster encoded on ICE*Sp*23FST81. This is predicted to produce a monomeric lantibiotic, most similar to lichenicidin and mersacidin, produced by *Bacillus* species. The known ring structure of mersacidin is indicated, again showing that the functional, ring-forming residues are largely conserved in this new lantibiotic. It is the third ring (the most strongly conserved between the three sequences) that is thought to be most important for the bactericidal effects of mersacidin, resulting from the inhibition of peptidoglycan transglycosylation.

until the ICE has saturated the available population of potential hosts is likely to facilitate its horizontal transfer.

ICEs have evidently played an important role in pneumococcal evolution. Given their wide potential host range, as exemplified by the detection of a conjugative element in *A. junii* apparently similar in structure to ICE*Sp*23FST81, and their ability to traffic long stretches of exogenous DNA, ICEs are likely to contribute a significant proportion of the *S. pneumoniae* pan-genome. The diverse "cargo" carried provides a large gene pool that can facilitate the rapid adaptation of the host species to new selective pressures, as illustrated by the case of antibiotic resistance in general, and chloramphenicol resistance in particular. Analogously, the presence of SOS response genes on pneumococcal conjugative elements may be protective against a different antimicrobial compound, hydrogen peroxide. At relatively low concentrations, this reactive oxygen species causes cell death through chromosomal damage (28); hence, a functional nucleotide excision repair pathway may increase the ability of the pneumococcus to tolerate H_2O_2 by improving the repair of oxidized DNA. However, the most significant change in the pneumococcal lineage to have been driven by a conjugative element may have been the insertion of the transposon that gave rise to PPI-1. The avirulence of mutants lacking genes in the 5′ region of PPI-1 in the mouse, as

well as the absence of this genome feature even in closely related streptococci, suggests that ICEs are also likely to have originally played an important role in the emergence of the pathogenic capability of pneumococci, as well as being important in their more recent adaptation to the challenge of antimicrobial chemotherapies.

Even in a species that exchanges genetic material as readily as *S. pneumoniae*, a small number of clones dominate the population of antibiotic-resistant pneumococci (36), suggesting that resistance must arise in a genetic background conducive to transmission for a multi-drug-resistant strain to emerge as a pandemic clone. Although much of the above analysis has focused on conjugative elements, factors elsewhere in the host chromosome, such as penicillin resistance, must also be important in aiding the spread of the strain, or else multidrug resistance would likely be distributed more quickly through conjugative transfer of ICE rather than clonal dissemination. It seems likely that pneumococcal pandemic clones represent chance combinations of horizontally acquired genetic elements that interact to allow the strain to outcompete other genotypes. The relative contributions of different segments of the chromosome to the emergence of such clones cannot be assessed by techniques such as multilocus sequence typing and BOX fingerprinting, which focus on regions of the genome that vary primarily through genetic drift. High-throughput sequencing and chromosome-wide genotyping technologies will permit a more precise knowledge of population structure, and provide more clues as to why certain lineages rise to international prevalence.

ACKNOWLEDGMENTS

We thank Ann Cronin, Claire Arrowsmith, Halina Norbertczak, Angela Lord, Suzanne Sanders, John Woodward, and the Pathogen Production Team at the Wellcome Trust Sanger Institute for their help with sequencing.

Work in Glasgow is supported by the BBSRC, MRC, and Wellcome Trust.

REFERENCES

1. **Ayoubi, P., A. O. Kilic, and M. N. Vijayakumar.** 1991. Tn*5253*, the pneumococcal omega (cat tet) BM6001 element, is a composite structure of two conjugative transposons, Tn*5251* and Tn*5252*. J. Bacteriol. **173:**1617–1622.
2. **Bagnoli, F., M. Moschioni, C. Donati, V. Dimitrovska, I. Ferlenghi, C. Facciotti, A. Muzzi, F. Giusti, C. Emolo, A. Sinisi, M. Hilleringmann, W. Pansegrau, S. Censini, R. Rappuoli, A. Covacci, V. Masignani, and M. A. Barocchi.** 2008. A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. J. Bacteriol. **190:**5480–5492.
3. **Barocchi, M. A., J. Ries, X. Zogaj, C. Hemsley, B. Albiger, A. Kanth, S. Dahlberg, J. Fernebro, M. Moschioni, V. Masignani, K. Hultenby, A. R. Taddei, K. Beiter, F. Wartha, A. von Euler, A. Covacci, D. W. Holden, S. Normark, R. Rappuoli, and B. Henriques-Normark.** 2006. A pneumococcal pilus influences virulence and host inflammatory responses. Proc. Natl. Acad. Sci. USA **103:**2857–2862.
4. **Bateman, A., M. T. G. Holden, and C. Yeats.** 2005. The G5 domain: a potential *N*-acetylglucosamine recognition domain involved in biofilm formation. Bioinformatics **21:**1301–1303.
5. **Bergmann, S., and S. Hammerschmidt.** 2006. Versatility of pneumococcal surface proteins. Microbiology **152:**295–303.
6. **Berriman, M., and K. Rutherford.** 2003. Viewing and annotating sequence data with Artemis. Brief. Bioinform. **4:**124–132.
7. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30:**157–166.
8. **Bogaert, D., R. de Groot, and P. W. M. Hermans.** 2004. *Streptococcus pneumoniae* colonization: the key to pneumococcal disease. Lancet Infect. Dis. **4**:144–154.
9. **Bogaert, D., A. van Belkum, M. Sluijter, A. Luijendijk, R. de Groot, H. C. Rumke, H. A. Verbrugh, and P. W. M. Hermans.** 2004. Colonization by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. Lancet **363:**1871–1872.
10. **Bonfield, J. K., K. F. Smith, and R. Staden.** 1995. A new DNA sequence assembly program. Nucleic Acids Res. **23:**4992–4999.
11. **Brown, J. S., S. M. Gilliland, and D. W. Holden.** 2001. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. Mol. Microbiol. **40:**572–585.
12. **Brown, J. S., S. M. Gilliland, B. G. Spratt, and D. W. Holden.** 2004. A locus contained within a variable region of pneumococcal pathogenicity island 1 contributes to virulence in mice. Infect. Immun. **72:**1587–1593.
13. **Brueggemann, A. B., D. T. Griffiths, E. Meats, T. Peto, D. W. Crook, and B. G. Spratt.** 2003. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. J. Infect. Dis. **187:**1424–1432.
14. **Burrus, V., G. Pavlovic, B. Decaris, and G. Guédon.** 2002. Conjugative transposons: the tip of the iceberg. Mol. Microbiol. **46:**601–610.
15. **Camilli, R., E. Pettini, M. D. Grosso, G. Pozzi, A. Pantosti, and M. R. Oggioni.** 2006. Zinc metalloproteinase genes in clinical isolates of *Streptococcus pneumoniae*: association of the full array with a clonal cluster comprising serotypes 8 and 11A. Microbiology **152:**313–321.
16. **Carver, T. J., K. M. Rutherford, M. Berriman, M.-A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423.
17. **Claverys, J. P., B. Martin, and L. S. Havarstein.** 2007. Competence-induced fratricide in streptococci. Mol. Microbiol. **64:**1423–1433.
18. **Coffey, T. J., C. G. Dowson, M. Daniels, J. Zhou, C. Martin, B. G. Spratt, and J. M. Musser.** 1991. Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of *Streptococcus pneumoniae*. Mol. Microbiol. **5:**2255–2260.
19. **Coffey, T. J., M. C. Enright, M. Daniels, J. K. Morona, R. Morona, W. Hryniewicz, J. C. Paton, and B. G. Spratt.** 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*. Mol. Microbiol. **27:**73–83.
20. **Coffey, T. J., M. C. Enright, M. Daniels, P. Wilkinson, S. Berreøn, A. Fenoll, and B. G. Spratt.** 1998. Serotype 19A variants of the Spanish serotype 23F multiresistant clone of *Streptococcus pneumoniae*. Microb. Drug Resist. **4:**51–55.
21. **Corso, A., E. P. Severina, V. F. Petruk, Y. R. Mauriz, and A. Tomasz.** 1998. Molecular characterization of penicillin-resistant *Streptococcus pneumoniae* isolates causing respiratory disease in the United States. Microb. Drug Resist. **4:**325–337.
22. **Crook, D. W., A. B. Brueggemann, K. L. Sleeman, and T. E. A. Peto.** 2004. Pneumococcal carriage. ASM Press, Washington, DC.
23. **Dawid, S., A. M. Roche, and J. N. Weiser.** 2007. The *blp* bacteriocins of *Streptococcus pneumoniae* mediate intraspecies competition both in vitro and in vivo. Infect. Immun. **75:**443–451.
24. **Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679.
25. **de Saizieu, A., C. Gardes, N. Flint, C. Wagner, M. Kamber, T. J. Mitchell, W. Keck, K. E. Amrein, and R. Lange.** 2000. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. J. Bacteriol. **182:**4696–4703.
26. **Dopazo, J., A. Mendoza, J. Herrero, F. Caldara, Y. Humbert, L. Friedli, M. Guerrier, E. Grand-Schenk, C. Gandin, M. de Francesco, A. Polissi, G. Buell, G. Feger, E. Garcia, M. Peitsch, and J. F. Garcia-Bustos.** 2001. Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate. Microb. Drug Resist. **7:**99–125.
27. **Erill, I., S. Campoy, and J. Barbe.** 2007. Aeons of distress: an evolutionary perspective on the bacterial SOS response. FEMS Microbiol. Rev. **31:**637–656.
28. **Fang, F. C.** 2004. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat. Rev. Microbiol. **2:**820–832.
29. **Forbes, M. L., E. Horsey, N. L. Hiller, F. J. Buchinsky, J. D. Hayes, J. M. Compliment, T. Hillman, S. Ezzo, K. Shen, R. Keefe, K. Barbadora, J. C. Post, F. Z. Hu, and G. D. Ehrlich.** 2008. Strain-specific virulence phenotypes of *Streptococcus pneumoniae* assessed using the *Chinchilla laniger* model of otitis media. PLoS ONE. **3:**e1969–e1969.
30. **Gindreau, E., R. Lopez, and P. Garcia.** 2000. MM1, a temperate bacteriophage of the type 23F Spanish/USA multiresistant epidemic clone of *Streptococcus pneumoniae*: structural analysis of the site-specific integration system. J. Virol. **74:**7803–7813.
31. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389–1406.
32. **Hill, P. C., A. Akisanya, K. Sankareh, Y. B. Cheung, M. Saaka, G. Lahai, B. M. Greenwood, and R. A. Adegbola.** 2006. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian villagers. Clin. Infect. Dis. **43:**673–679.
33. **Hiller, N. L., B. Janto, J. S. Hogg, R. Boissy, S. Yu, E. Powell, R. Keefe, N. E. Ehrlich, K. Shen, J. Hayes, K. Barbadora, W. Klimke, D. Dernovoy, T. Tatusova, J. Parkhill, S. D. Bentley, J. C. Post, G. D. Ehrlich, and F. Z. Hu.** 2007. Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. J. Bacteriol. **189:**8186–8195.

34. **Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass.** 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709–5717.
35. **Jacobs, M. R., C. E. Good, B. Beall, S. Bajaksouzian, A. R. Windau, and C. G. Whitney.** 2008. Changes in serotypes and antimicrobial susceptibility of invasive *Streptococcus pneumoniae* strains in Cleveland: a quarter century of experience. J. Clin. Microbiol. **46:**982–990.
36. **Klugman, K. P.** 2002. The successful clone: the vector of dissemination of resistance in *Streptococcus pneumoniae*. J. Antimicrob. Chemother. **50:**1–5.
37. **Kyaw, M. H., R. Lynfield, W. Schaffner, A. S. Craig, J. Hadler, A. Reingold, A. R. Thomas, L. H. Harrison, N. M. Bennett, M. M. Farley, R. R. Facklam, J. H. Jorgensen, J. Besser, E. R. Zell, A. Schuchat, C. G. Whitney, et al.** 2006. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant *Streptococcus pneumoniae*. N. Engl. J. Med. **354**:1455–1463.
38. **Lanie, J. A., W.-L. Ng, K. M. Kazmierczak, T. M. Andrzejewski, T. M. Davidsen, K. J. Wayne, H. Tettelin, J. I. Glass, and M. E. Winkler.** 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. J. Bacteriol. **189:**38–51.
39. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. CLUSTAL W and CLUSTAL X version 2.0. Bioinformatics **23:**2947–2948.
40. **Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison, and D. W. Holden.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:** 555–571.
41. **Loeffler, J. M., and V. A. Fischetti.** 2006. Lysogeny of *Streptococcus pneumoniae* with MM1 phage: improved adherence and other phenotypic changes. Infect. Immun. **74:**4486–4495.
42. **Lux, T., M. Nuhn, R. Hakenbeck, and P. Reichmann.** 2007. Diversity of bacteriocins and activity spectrum in *Streptococcus pneumoniae*. J. Bacteriol. **189:**7741–7751.
43. **Martin, B., O. Humbert, M. Camara, E. Guenzi, J. Walker, T. Mitchell, P. Andrew, M. Prudhomme, G. Alloing, R. Hakenbeck, D. A. Morrison, G. J. Boulnois, and J.-P. Claverys.** 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. Nucleic Acids Res. **20:**3479–3483.
44. **McCool, T. L., T. R. Cate, G. Moy, and J. N. Weiser.** 2002. The immune response to pneumococcal proteins during experimental human carriage. J. Exp. Med. **195:**359–365.
45. **Mingoia, M., M. Vecchi, I. Cochetti, E. Tili, L. A. Vitali, A. Manzin, P. E. Varaldo, and M. P. Montanari.** 2007. Composite structure of *Streptococcus pneumoniae* containing the erythromycin efflux resistance gene *mef*(I) and the chloramphenicol resistance gene *catQ*. Antimicrob. Agents Chemother. **51:**3983–3987.
46. **Møller, P., and H. Wallin.** 1998. Adduct formation, mutagenesis and nucleotide excision repair of DNA damage produced by reactive oxygen species and lipid peroxidation product. Mutat. Res. **410:**271–290.
47. **Muñoz, R., T. J. Coffey, M. Daniels, C. G. Dowson, G. Laible, J. Casal, R. Hakenbeck, M. Jacobs, J. M. Musser, and B. G. Spratt.** 1991. Intercontinental spread of a multiresistant clone of serotype 23F *Streptococcus pneumoniae*. J. Infect. Dis. **164:**302–306.
48. **Munoz-Najar, U., and M. N. Vijayakumar.** 1999. An operon that confers UV resistance by evoking the SOS mutagenic response in streptococcal conjugative transposon Tn*5252*. J. Bacteriol. **181:**2782–2788.
49. **Obert, C., J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. I. Tuomanen, and C. J. Orihuela.** 2006. Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect. Immun. **74:**4766–4777.
50. **Obregón, V., J. L. García, E. García, R. López, and P. García.** 2003. Genome organization and molecular analysis of the temperate bacteriophage MM1 of *Streptococcus pneumoniae*. J. Bacteriol. **185:**2362–2368.
51. **Oggioni, M. R., and J. P. Claverys.** 1999. Repeated extragenic sequences in prokaryotic genomes: a proposal for the origin and dynamics of the RUP element in *Streptococcus pneumoniae*. Microbiology **145:**2647–2653.
52. **Oggioni, M. R., G. Memmi, T. Maggi, D. Chiavolini, F. Iannelli, and G. Pozzi.** 2003. Pneumococcal zinc metalloproteinase ZmpC cleaves human matrix metalloproteinase 9 and is a virulence factor in experimental pneumonia. Mol. Microbiol. **49:**795–805.
53. **Ojo, K. K., N. L. Ruehlen, N. S. Close, H. Luis, M. Bernardo, J. Leitao, and M. C. Roberts.** 2006. The presence of a conjugative gram-positive Tn*2009* in gram-negative commensal bacteria. J. Antimicrob. Chemother. **57:**1065–1069.
54. **Orihuela, C. J., J. N. Radin, J. E. Sublett, G. Gao, D. Kaushal, and E. I. Tuomanen.** 2004. Microarray analysis of pneumococcal gene expression during invasive disease. Infect. Immun. **72:**5582–5596.
55. **Parkhill, J., M. Achtman, K. D. James, S. D. Bentley, C. Churcher, S. R. Klee, G. Morelli, D. Basham, D. Brown, T. Chillingworth, R. M. Davies, P. Davis, K. Devlin, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, S. Moule, K. Mungall, M. A. Quail, M. A. Rajandream, K. M. Rutherford, M. Simmonds, J. Skelton, S. Whitehead, B. G. Spratt, and B. G. Barrell.** 2000. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. Nature **404:**502–506.
56. **Paterson, G. K., L. Nieminen, J. M. C. Jefferies, and T. J. Mitchell.** 2008. PclA, a pneumococcal collagen-like protein with selected strain distribution, contributes to adherence and invasion of host cells. FEMS Microbiol. Lett. **285**:170–176.
57. **Pericone, C. D., K. Overweg, P. W. M. Hermans, and J. N. Weiser.** 2000. Inhibitory and bactericidal effects of hydrogen peroxide production by *Streptococcus pneumoniae* on other inhabitants of the upper respiratory tract. Infect. Immun. **68:**3990–3997.
58. **Pletz, M. W. R., L. McGee, J. Jorgensen, B. Beall, R. R. Facklam, C. G. Whitney, and K. P. Klugman.** 2004. Levofloxacin-resistant invasive *Streptococcus pneumoniae* in the United States: evidence for clonal spread and the impact of conjugate pneumococcal vaccine. Antimicrob. Agents Chemother. **48:**3491–3497.
59. **Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari, and D. Simon.** 1998. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:**5620–5629.
60. **Reinert, R. R., A. Ringelstein, M. van der Linden, M. Y. Cil, A. Al-Lahham, and F.-J. Schmitz.** 2005. Molecular epidemiology of macrolide-resistant *Streptococcus pneumoniae* isolates in Europe. J. Clin. Microbiol. **43:**1294–1300.
61. **Rose, L., P. Shivshankar, E. Hinojosa, A. Rodriguez, C. J. Sanchez, and C. J. Orihuela.** 2008. Antibodies against PsrP, a novel *Streptococcus pneumoniae* adhesin, block adhesion and protect mice against pneumococcal challenge. J. Infect. Dis. **198:**375–383.
62. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
63. **Sandgren, A., K. Sjostrom, B. Olsson-Liljequist, B. Christensson, A. Samuelsson, G. Kronvall, and B. Henriques-Normark.** 2004. Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*. J. Infect. Dis. **189:**785–796.
64. **Scott, J. A. G.** 2007. The preventable burden of pneumococcal disease in the developing world. Vaccine **25:**2398–2405.
65. **Shen, K., J. Gladitz, P. Antalis, B. Dice, B. Janto, R. Keefe, J. Hayes, A. Ahmed, R. Dopico, N. Ehrlich, J. Jocz, L. Kropp, S. Yu, L. Nistico, D. P. Greenberg, K. Barbadora, R. A. Preston, J. C. Post, G. D. Ehrlich, and F. Z. Hu.** 2006. Characterization, distribution, and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*. Infect. Immun. **74:**321–330.
66. **Singleton, R. J., T. W. Hennessy, L. R. Bulkow, L. L. Hammitt, T. Zulz, D. A. Hurlburt, J. C. Butler, K. Rudolph, and A. Parkinson.** 2007. Invasive pneumococcal disease caused by nonvaccine serotypes among Alaska native children with high levels of 7-valent pneumococcal conjugate vaccine coverage. JAMA **297:**1784–1792.
67. **Sjostrom, K., C. Spindler, A. Ortqvist, M. Kalin, A. Sandgren, S. Kuhlmann-Berenzon, and B. Henriques-Normark.** 2006. Clonal and capsular types decide whether pneumococci will act as a primary or opportunistic pathogen. Clin. Infect. Dis. **42:**451–459.
68. **Slamti, L., and D. Lereclus.** 2002. A cell-cell signaling peptide activates the PlcR virulence regulon in bacteria of the *Bacillus cereus* group. EMBO J. **21:**4550–4559.
69. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:**1596–1599.
70. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498–506.
71. **Vernikos, G. S., and J. Parkhill.** 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. Bioinformatics **22:**2196–2203.
72. **Wagner, P. L., and M. K. Waldor.** 2002. Bacteriophage control of bacterial virulence. Infect. Immun. **70:**3985–3993.
73. **Widdowson, C. A., P. V. Adrian, and K. P. Klugman.** 2000. Acquisition of chloramphenicol resistance by the linearization and integration of the entire staphylococcal plasmid pC194 into the chromosome of *Streptococcus pneumoniae*. Antimicrob. Agents Chemother. **44:**393–395.
74. **Willey, J. M., and W. A. van der Donk.** 2007. Lantibiotics: peptides of diverse structure and function. Annu. Rev. Microbiol. **61:**477–501.